DOCUMENT RESUME

ED 041 947 TM 000 039

AUTHOR	Eash, Maurice J.
TITLE	Developing An Instrument for the Assessment of Instructional Materials (Form IV).
INSTITUTION	City Univ. of New York, N.Y. Hunter Coll.
SPONS AGENCY	Bureau of Education for the Handicapped (DHEW/OE), Washington, D.C.
BUREAU NO	BR-7-0729
PUB DATE	Mar 70
GRANT	OEG-1-2-070729-3104(032)
NOTE	37p.; Paper presented at the annual meeting of the American Educational Research Association, Minneapolis, Minn., March 1970
EDRS PRICE	EDRS Price MF-$0.25 HC-$1.95
DESCRIPTORS	Evaluation, *Evaluation Criteria, Evaluation Techniques, *Instructional Design, *Instructional Materials, *Measurement Instruments, Measurement Techniques, Rating Scales, *Textbook Evaluation, Textbooks

ABSTRACT

This instrument is designed to provide data on the characteristics of the instructional design of classroom materials. Data is gathered on four constructs: objectives, organization of the material (scope and sequence), methodology, and evaluation. Questions elicit a mixture of yes-no, check list, and summary rating scale answers with open ended responses permitted where an exception to the listed approaches is encountered. The rater progresses through an atomistic analysis of the material to a numerical summary rating of the construct. In this study twenty-five graduate students rated two instructional packages, a reading package and a science package. Ratings were done individually and by groups. Individual responses showed agreement in 56 of 64 responses. Some of the limitations of the questions and the few disparities between individual and team evaluations are discussed. Inter-rater reliability scores both for each of the ratings on the constructs and the overall score were greater than .9. Inter-item reliability estimates were: objectives .38, organization of material .37, methodology .77, evaluation .99, and overall .55. A follow-up study with teachers after they have assessed materials with the instrument is suggested. Finally, general issues in instructional design are discussed. The instrument and a comprehensive glcssary are included. (Author/GS)

ED041947

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

DEVELOPING AN INSTRUMENT FOR THE ASSESSMENT OF INSTRUCTIONAL MATERIALS (FORM IV)[1]

MAURICE J. EASH, UNIVERSITY OF ILLINOIS (CHICAGO CIRCLE)

The most common description applied to the salient characteristic of education in the past decades has been growth. Pupil population, faculty, buildings, auxiliary personnel, and instructional materials have all reflected the growth phenomenon. While the upshot of the expansion effort has created opportunities and stimulated experimental and developmental activity, the magnitude of the process in many areas has been accompanied by attendant problems. The instrument which is described in this paper is an outgrowth of an attempt to deal with the problem of the embarrassment of riches in instructional materials that are available to schools; the necessity to select from this cornucopia of materials and to implement them effectively in the classroom. For not only has the range of materials increased in this decade, but the sophistication of the instructional design has advanced to the stage where change of instructional materials frequently entails extensive retraining of teachers. Moreover, failure to implement instructional materials in the classroom in keeping with the instructional design requirements of the originator and producer has come to be accepted as a frequent cause for failure for instructional innovations once the materials are outside the developmental and/or experimental setting.

This instrument then was developed to assist in providing data on two broad questions: 1) What materials shall we select for use in the classroom? 2) What are the characteristics of the instructional design of the material and what will it take to implement it effectively given

---

[1] Prepared for the American Educational Research Association Annual Conference, Division B, Minneapolis, Minnesota, March 1970.

TM 000 039

the demands of a particular program? It is important to make doubly clear at this point that the instrument does not make the judgment, but it does systematize the gathering of data, serves to focus the analysis, and provides a data base for a critical judgment which has the further advantage of being grounded in the literature and research on instruction. Thus, hopefully, the decision-making process in the area of choosing instructional materials will be removed from the realm of vague intuitions to one where the selected materials will have a degree of predictive validity for learning outcomes with students. That work on the problem of instrumentation may have made some initial progress along the continuum toward rational decision-making in choice of instructional materials is the subject of the remainder of this paper.

This paper reports the findings of a field test of Form IV of an instrument developed for assessing instructional materials and some general formative observations on the instrument garnered from use in a number of field trials. Before launching into the presentation of the findings, a brief description of the instrument seems in order.

As designed, the instrument serves to illuminate the instructional design potentialities of a range of instructional materials through eliciting of data on essential constructs generally accepted as central to micro and macro designs of curriculum. In this instrument, these constructs are labeled: 1) objectives, 2) organization of the material (scope and sequence), 3) methodology and, 4) evaluation. Under each of these major constructs are listed as many of the customary approaches used in meeting these constructs in instructional materials as was feasible without overextending the length of the instrument. An open ended response is permitted where an exception to the listed approaches

is encountered. For the purpose of developing a summary statement on the material, the rater is further encouraged to note examples of how the construct is being satisfied. In addition to this decomposition analysis which extracts and selects data on how the instructional materials are satisfying the four constructs of an instructional design, there are other items which seek information on the development of the materials, primarily whether the materials were field tested or researched with prospective consumers. At the end of the section on each construct, and at the end of the instrument there is a seven point scale in which the rater is asked to render a judgment on each of the four constructs, and a comprehensive judgment on the overall worth of the instructional material. These scales have three points defined through descriptions of characteristics of materials which would fall at these points; materials having a mix of these characteristics presumably would fall somewhere in between these points. As outlined, the basic scheme is to have the rater progress through an atomistic analysis of the material then move to a summary rating of the construct, a process devised after bitter experience with wildly inconsistent ratings when only the summary qualitative ratings were used to judge the constructs of the materials. Upon completion of this assessment, the rater is asked to prepare a short summary statement, now presumably arrived at through a more competent understanding of the materials and their potential in an instructional setting.

II Procedures

The subjects for this study were twenty-five graduate students in an advanced class in curriculum design. Of the twenty-five subjects, eleven were elementary teachers, fourteen were secondary teachers and all

had over five years teaching experience. Two types of materials were used with the instrument. In the first trial a sixth grade reading package from a widely used reading series composed of a teacher's manual, the student's reader and workbook were assessed individually by the twenty-five subjects. Subsequently the twenty-five subjects were formed into seven teams for a second trial and collective ratings on the reading package were rendered.[2] The third trial of the instrument was conducted on a curriculum bulletin in science, grade seven, developed by a major city system.[3] Prepared with the expressed purpose of assisting the classroom teacher in science through providing a comprehensive micro design for twenty-four classroom and laboratory lessons, the bulletin was one of a series and featured the chemistry of matter: elements, compounds and mixtures, and atomic theory. Only judgments by the seven teams were gathered on the science curriculum bulletin.

In the first trial with the reading package subjects had been asked to fill out the forms individually maintaining a record of the time they had spent. They reported a range of time required from two to four hours with a mean of three hours. The collective team judgments in the two trials were gathered in a two hour period. After becoming familiar with the instrument, subjects could individually assess material with rapidity, and the team sessions time was mainly devoted to reconciliation of the disparate judgments. While the reading package was more complex and of greater length than the chemistry unit, the subjects were, with the exception of four individuals, who were science teachers, not as familiar with the content of the chemistry unit and the elements of length of reading package and unfamiliarity with the science content balanced out. The two hour time units appeared to be sufficient time for the groups to complete their collective assessments of the learning packages.

III Results

Following completion of the three trials, the data were compiled by item, and percentage of individuals' response computed for each item of the assessment. The frequency of response for each item was also computed. Since the purposes of this study are primarily formative and the data are presented as illustrative, in the interest of brevity only the individual item numeration for reading is presented.

TABLE I

ITEM TABULATION OF RESPONSES BY INDIVIDUALS (25)
TEAMS (7) ON THE LEARNING PACKAGE IN READING

QUESTION CATEGORY	PERCENTAGE FOR 25 INDIVIDUALS					FREQUENCIES FOR 7 TEAMS		
	YES	NO	OMIT	MEAN	S.D.	YES	NO	OMIT
I. OBJECTIVES								
A.	92	0	8	1	0	7	0	0
1.	100	0	0	1	0	6	0	1
2.	96	0	0	1	0	6	0	1
3.	92	8	0	.92	.27	6	0	1
4. Instructions								
a.	92	4	4	.96	.20	6	0	1
b.	68	28	4	.71	.45	4	2	1
c.	20	76	4	.21	.41	1	5	1
5. Anecdotal								
B.	12	12	76	.50	.50	0	0	7
1. Anecdotal								
C. Instruction								
1.	100	0	0	1	0	6	0	1
2.	96	4	0	.96	.20	6	0	1
3.	92	8	0	.92	.27	6	0	1
4. Instructions								
a.	92	0	8	1	0	7	0	0
b.	32	0	68	1	0	1	0	6
c.	100	0	0	1	0	7	0	0
d.	96	0	4	1	0	7	0	0
5. Instruction								
a.	84	0	16	1	0	5	0	2
b.	92	0	8	1	0	7	0	0
c.	100	0	0	1	0	7	0	0
d.	84	0	16	1	0	6	0	1
D. (See Tables II and III)								

TABLE I - continued

	YES	NO	OMIT	MEAN	S.D.	YES	NO	OMIT
II. ORGANIZATION OF THE MATERIAL								
A.	84	16	0	.84	.37	6	1	0
B. Instructions								
1.	4	0	96	1	0	0	0	7
2.	88	0	12	1	0	5	0	2
3.	20	0	80	1	0	1	0	6
4.	68	0	32	1	0	5	0	2
5.	64	0	36	1	0	4	0	3
6.	16	0	84	1	0	0	0	7
C. Anecdotal								
D.	84	0	16	1	0	7	0	0
1. Instructions								
a.	92	4	4	.96	.20	6	1	0
b.	28	52	20	.35	.48	2	4	1
c.	100	0	0	1	0	7	0	0
d.	92	0	8	1	0	7	0	0
e. Anecdotal								
2. Instructions								
a.	64	12	24	.84	.36	4	3	0
b.	56	16	28	.78	.42	3	3	1
E.								
1. Instructions								
a.	80	0	20	1	0	5	0	2
b.	28	0	72	1	0	0	0	7
c.	52	0	48	1	0	2	0	5
d.	80	0	20	1	0	5	0	2
e. Anecdotal								
F. Anecdotal								
G. (See Tables II and III)								
III. METHODOLOGY								
A.	100	0	0	1	0	7	0	0
B.	96	0	4	1	0	7	0	0
1. Instructions								
a.	88	12	0	.88	.32	7	0	0
b.	20	68	12	.23	.42	0	5	2
c.	60	20	20	.75	.43	3	2	2
d.	60	20	20	.75	.43	5	0	2
e.	76	12	12	.86	.34	6	1	0
f.	84	16	0	.86	.34	6	1	0
g.	52	44	4	.54	.50	4	2	1
C.	72	20	8	.22	.41	0	6	1
1. (See Tables II and III)								
2.	4	96	0	.04	.20	1	6	0
3.	0	96	4	.00	.00	0	7	0
4.[1]								
a. Anecdotal								
b. Anecdotal								

TABLE I - continued

	YES	NO	OMIT	MEAN	S.D.	YES	NO	OMIT
D. Anecdotal								
E. (See Tables II and III)								
IV. EVALUATION								
A.	88	8	4	.92	.28	6	0	1
1. Instructions								
a.	92	0	8	1	0	7	0	0
b.	96	0	4	1	0	7	0	0
c.	4	0	96	1	0	0	0	7
d.	52	0	48	1	0	3	0	4
2.	92	0	8	1	0	7	0	0
3. Instructions								
a.	76	0	24	1	0	5	0	2
b.	72	0	28	1	0	5	0	2
c.	84	0	16	1	0	6	0	1
d.	76	0	24	1	0	5	0	2
e.	70	0	24	1	0	5	0	2
B.	96	4	0	.96	.20	7	0	0
C.	56	32	12	.64	.48	3	3	1
D.	0	96	4	0	0	0	7	0
E. Anecdotal								
F. (See Tables II and III)								
V. COMMENT								
A. Anecdotal								
B. (See Tables II and III)								

[1]Responses are: most succeeded 0; approximately half succeeded 20%; few succeeded 4%; omitted 76%.

Tables II, III, and IV present the quantitative ratings assigned on the seven point scale to each of the four constructs and the overall assessment rating of the learning packages under section V, Comments, by individuals and the two teams.

TABLE II
SUMMARY RATINGS OF CONSTRUCTS ON SEVEN POINT SCALE (READING)

CONSTRUCTS	FREQUENCIES FOR 25 INDIVIDUALS								
	1	2	3	4	5	6	7	MEAN	S.D.
Objectives I	0	0	1	6	12	4	2	5.00	0.95
Organization II	0	1	1	11	8	3	1	4.56	1.04
Methodology III	0	0	3	12	5	4	1	4.52	1.04
Evaluation IV	0	3	2	14	5		1	4.00	1.07
Overall Assessment V	0	0	2	8	12	3	0	4.64	0.80

TABLE III
SUMMARY RATINGS BY TEAMS OF CONSTRUCTS ON SEVEN POINT SCALE (READING)

CONSTRUCTS	FREQUENCIES FOR 7 TEAMS								
	1	2	3	4	5	6	7	MEAN	S.D.
Objectives I	0	0	0	2	4	0	1	5.00	1.00
Organizations II	0	0	1	2	3	1	0	4.57	0.97
Methodology III	1	0	1	3	2	0	0	3.71	1.37
Evaluation IV	0	1	0	5	1	0	0	3.85	0.90
Overall Assessment V	0	0	1	3	3	0	0	4.28	0.75

TABLE IV
SUMMARY RATINGS BY TEAMS OF CONSTRUCTS ON SEVEN POINT SCALE (SCIENCE)

CONSTRUCTS	1	2	3	4	5	6	7	MEAN	S.D.
Objectives I	0	1	4	1	1	0	0	3.28	0.94
Organization II	0	0	1	4	1	1	0	4.28	0.94
Methodology III	0	0	1	3	2	1	0	4.42	0.97
Evaluation IV	0	3	2	1	0	1	0	3.14	1.46
Overall Assessment V	0	0	2	3	2	0	0	4.00	0.81

Previously it has been pointed out that the instrument elicits considerable anecdotal data. These are not reported except for those from the final section where the raters sum up the strengths and weaknesses of the materials. Table V presents the summaries of the strengths and weaknesses of the learning package on reading as prepared under section V, Comment, of the instrument. The statements in Table V have been drawn from the individual ratings. For purposes of analysis they have further been divided into elementary and secondary respondents. Tables VI and VII present the summary statements on the strengths and weaknesses of the two learning packages on reading and science as recorded by the teams.

TABLE V

SUMMARY COMMENTS ON OVERALL ASSESSMENT BY INDIVIDUALS
OF LEARNING PACKAGE (READING)

ELEMENTARY TEACHERS

Strengths:

1. Topics are very diverse and appeal to the children.
2. Pictures are colorful; print is large.
3. Tests in workbook after each unit.
4. Exercises on skills to be developed appear in T.G.
5. Suggestions on approaches to each story.
6. Basic reading test is available.
7. Workbook is good for independent activities.
8. Stories selected have high interest level.
9. T.G. presents a variety of methodologies.
10. Good supplementary reading list presented, keeps kids in mind.
11. Variety of subject matter geared to this age group (6th grade).
12. Objectives good, varied, and clearly stated.
13. Provides for evaluation, feedback, and reinforcement.
14. Will adapt easily to higher ability levels (workbook).
15. Has criterion reference procedures for evaluation.
16. Provides norm references.
17. Good objectives (behavioral).

TABLE V - continued

Weaknesses:

1. No level of performance in behavioral objectives.
2. No immediate evaluative feedback for student.
3. No indication that material has been field tested.
4. Psychomotor skills are not clearly indicated.
5. Strongly teacher-centric.
6. Evaluation poor; methodology poor; no allowance for individual differences.
7. Book itself does not provide materials for the below or above average.
8. I question the interest level of the selections.
9. Criteria for evaluation not clearly stated.
10. No objectives for children set forth in workbook exercises: why they are doing the exercises.
11. Not enough reinforcement of skills in a logical sequence.
12. Not enough variety of activities in workbook.
13. Additional material for evaluation of concepts is needed.

SECONDARY TEACHERS

Strengths:

1. Strength lies in its objectives and somewhat in scope and sequence.
2. Topics selected have a high interest level and are timely.
3. Graphic presentations are excellent.
4. Lessons help develop cognitive and subject skills.
5. Includes many aids for the teacher.
6. A thoroughly linguistic base.
7. Stimulates extensive reading.
8. Has workbook.

TABLE V - continued

9. Survey test and basic reading test available.
10. Good modes of transaction and evaluation in reading skills area.
11. Valuable for the new teacher who needs a well structured package.
12. Logical organization.
13. Multi-ethnic appeal.
14. Remedial exercises to strengthen reading abilities.
15. Reading level appropriate for grade and pre-adolescent.
16. Allows teacher lots of variety in approach.

Weaknesses:

1. Range of evaluation limited to behavior product as criterion measure.
2. Provisions for measure of learner behavior should be included.
3. Evaluation limited to product with little process attention.
4. No immediate student feedback.
5. Does not appear to have been widely field tested.
6. Objectives do not specify level of performance. (Sometimes implied.)
7. Psycho-motor skills not enumerated in specific terms.
8. Perhaps the objectives could have been more clearly developed.
9. Very structured and does not permit for a variety of modes of transaction.
10. Weak in methodology and evaluation.
11. Over emphasized facts and knowledge in modes of transaction.
12. Overly teacher-centric.
13. Lacks reinforcement throughout.
14. Makes no provision for variations in individual ability.
15. Selection of content lacked imagination.
16. Teacher would need some training before using the reading materials.

TABLE VI
SUMMARY COMMENTS BY TEAMS ON
OVERALL ASSESSMENT OF LEARNING PACKAGE (READING)

Strengths:

1. It is structurally designed.
2. Scope and sequence well organized.
3. Variety of modes of transaction (especially in follow-up activities).
4. Criterion reference for evaluation - workbook.
   Norm reference for evaluation - test kit.
5. Materials highly structured.
6. Cognitive skills and objectives state the type of expected behavior and have structured formal and informal evaluation.
7. Objectives frequently behaviorally stated.
8. Multi-ethnic wide story appeal.
9. Could be useful to new teachers who need a structured program.
10. Availability of tests.
11. Suggested methodology as well as organization of material.
12. Stories selected for high interest level.
13. Various aids for the teacher in manual and workbook.
14. Methodology is spelled out for those who need it.
15. Good independent activities in workbook.
16. Tests in workbook for each unit also survey and inventory tests available.

Weaknesses:

1. Not enough variety in evaluation and reinforcement exercises for weaknesses.
2. More pupil centered activities needed.
3. Makes only limited provision for individual variations in ability and skills.

TABLE VI - continued

4. Teachers need some training to use overall program.
5. Graphic presentations poor.
6. Very teacher centric.
7. Overemphasis on facts, especially in the dominant modes of transaction.
8. Overly concerned with product.
9. Overly teacher centric approach.
10. Teacher centric to an extreme.
11. Little allowance for variety of modes of transaction - limited provision for individual differences.
12. Has not been evaluated as an effective teaching method.
13. Has a traditional subject-logic, teacher centric mode of transaction.
14. Levels of performance for objectives on a day to day basis not stated.
15. Lacks immediate feedback for pupils.
16. Modes of transaction and evaluation of attitudes are weak.
17. Psychomotor skills not enumerated in specific terms.
18. No evidence of field testing.
19. Lack of immediate feedback for students to get assistance.
20. No indication of field evaluation.

TABLE VII
SUMMARY STATEMENTS BY TEAMS ON
OVERALL ASSESSMENT OF LEARNING PACKAGE (SCIENCE)

Strengths:

1. Operational for the teacher without science background.
2. General objectives are stated.
3. Scope and sequence are set forth.
4. Scope and sequence are logical, organization is orderly and clear.
5. Part of a sequential K-12 program, clear on how it fits in.
6. Good format, understandable.
7. Some strengths on teacher orientation to the area.

Weaknesses:

1. No behavioral objectives.
2. Limited interrelationships between units.
3. Gap between goals and design, for example, little or no emphasis on process or discovery.
4. Few evaluation tools, no field testing.
5. Organization rigid and subject demand oriented.
6. Objectives not stated in behavioral terms, instructional objectives weak and difficult to operationalize.
7. Methodology does not provide for individual differences - but average in this respect.
8. Too factual oriented.
9. Teacher centric modes of transaction.
10. Poor evaluation tools in the package.
11. Scope and sequence lacks any behavioral statements.
12. Subject centered, product rather than process oriented.

IV DISCUSSION

Examination of the individual responses in Table I shows considerable agreement in 56 of the 64 responses. The greatest disagreement occurred on the following items.

I, A, 4, b.	Conditions under which it (objective) will appear.
II, B, 5.	Logical order - as a basis to organize materials.
II, D, 1, b.	Basis for scope and sequence: to a motor skill development.
II, D, 2, a.	Scope and sequence has been analyzed for appropriateness to students.
II, D, 2, b.	Scope and sequence have been analyzed for relationship to other material.
II, G.	Quantitative rating of organization of materials.
III, B, 1, c.	Does the mode of transaction require active student participation?
III, B, 1, g.	Does the mode of transaction: provide for variation among students - approaches to method?
IV, A, 1, d.	Do evaluation procedures emphasize affective responses?
IV, C.	Evaluation plan gives attention to both product and process.
IV, F.	Evaluation: quantitative rating.

In a few instances, disagreements among individuals were compressed and disappeared in the team judgments as in the example: I, B, "if there are no objectives stated for the use of the materials are the objectives instead implicit or readily obvious?" The weight of opinion among individuals where 76% had omitted the item carried over into the team responses and 100% of the teams omitted the item in a triumph of group pressure over the individual.

A somewhat comparable situation occurs where the team assessment did not correct individual errors in the objective section in item I,

A, 4, c, which asks, if objectives are stated in behavioral terms do they specify levels of performance. In individual's assessments five subjects said yes, nineteen said no, one omitted a response. One team said yes, five teams said no, one team omitted. In this section, raters asked to list examples of objectives and inspection of tne anecdotal comments would lend weight to the correctness of the majorities interpretation, since none of the objectives listed specify performance criteria. Examples of objectives listed are: "Teach pupils to listen, speak and write effectively and well," "Increase competence in reading skills and encourage personal reading," "Children use guide words to locate entries," "Children generalize about a main idea." At the center of this problem appears to have been the application of a definition to specific cases, and in the case of the team who responded affirmatively to the item it was the vagaries of chance clustering three subjects together who had made an incorrect assessment in their individual ratings.

On closer inspection, some of the disagreements are over items where examination of the material would seem to establish it as a fact and a respondence should readily ascertain whether it should be a yes or no response. As on Item II, D, 2, b, "Scope and sequence have been analyzed for relationship to other materials," three teams answered yes, three answered no. Evidently what was simply a straight forward question to the author posed an ambiguity for the raters. Considering the attention given to instructional design in a reading package, lack of unanimity in this item is troubling, and one suspects that scope and sequence was not seen as a unitary concept by the teams and served as a dual stimuli in the assessment, i.e., teams may have been answering two different questions.

Other problems resulting from differing subject've interpretation seem to have arisen in items II, D, 1, b, III, B, 1, c, and III, B, 1, g. Since there is a multitude of stimuli to judge in the reading package, literally hundreds of pupil activities, individuals and teams could very well have been using different samples of data to form their judgments.

Some additional paradoxes emerge when the statements by individuals on strengths and weaknesses are compared. (Table V) The same item occasionally appearing in both categories. Again the range of the stimuli to which the subjects were reacting may be a source of difficulty - as well as the problem of honest differences existing in subjective judgments as whether the stories in the reading package are of interest to students. It is of interest to note that a number of these paradoxes are ironed out in the teams' listing and strengths and weaknesses are more mutually exclusive in Table VI than in Table V. The hypothesis that resolution of the differences of individuals assessment comes through agreement on what stimuli to judge, is seen in the teams' judgment of the science bulletin in Table VII. Again the categories of strengths and weaknesses are practically mutually exclusive which greatly increases their value in rendering an overall summary judgment of the potential of the learning package for classroom use.

The inter-rater reliability, estimated by comparing the scores of every other subject (odd-even), indicates that for each of the ratings on the constructs and the overall score, the estimates showed greater than .9. On the other hand, inter-item reliability estimates, calculated for each subscore and for the overall score, were as follows: objectives .38, organization of material .37, methodology .77,

evaluation .99 (though this last statistic is suspect since four responses were inadvertantly omitted from the calculation), and overall .55. In this latter analysis although certain of the subscores suggest some interval consistency, in general the instrument on this factor had low reliability in this administration.

On the quantitative ratings of the two learning packages, Tables II, III and IV, there are only small differences in the mean ratings assigned to the constructs by individuals and the teams. However, the distribution of the ratings become slightly more compressed in the team ratings as can be seen in comparing Tables II, III and IV, noting limited numbers of ratings in the frequency columns of six and seven in the team ratings, and the generally narrower S.D. for the teams. There is a considerable degree of consistency in the average ratings assigned in these three trials from construct to construct and the means cluster around the midpoint of the scale.

While some experts have agreed that the categories of the instrument contain a high degree of content validity no criterion related validity studies have been made. A follow-up study on the use of instructional packages by teachers after they have assessed the materials with the instrument would provide useful data on the instrument's effectiveness in improving implementation as well as the value of its a priori evaluation of the learning packages.

Criticism has been voiced that the seven point scale at the end of each construct and for overall assessment requires judgments on a variety of stimuli and hence confuses the rater. Granted a rater must weigh the several factors in his mind as he makes a general judgment on the construct and overall worth of the material but such are the

decisions we make in most areas of life. Seldom are the variables discrete and free of ambiguity, and instructional packages being the organized complexities they are do pose similar problems eventuating in selections that are based upon compromise. Should we accept weak evaluation techniques and limited teacher aids on methodology for material that has carefully delineated objectives and a well prepared scope and sequence? Do we sacrifice poorly stated objectives for well developed methodology of proven interest to students? These are not atypical questions that face decision makers in materials selection. The instrument encourages examining the trade offs that must be made prior to making a judgment through helping a rater establish specifically what the options are in the design of the material.

Another issue that has been raised and grappled with is the feasibility of an instrument that stresses instructional design and ignores content. As the form is now constructed instructional materials that lean toward a programmed approach and stress instructional design over content would be favored in assessment. Repeated field trials with the instrument have found that instructional materials where design has been stressed such as in reading, and materials that have a tight internal logic as in arithmetic do lend themselves to easier assessment within the framework of the instrument. Nevertheless, they do not receive exceptionally high ratings since raters judge them against similar competing materials and not against materials in other subject fields where design has not been as prominent a concern. In early forms of the instrument, the issue of assessing content was examined, but other than the items on organization of materials (scope and sequence) the present instrument does not address this problem directly.

Even then, section II of the instrument does reflect the knotiness of the problem in its somewhat lower reliability and consistency. So far I have concluded that the judgment of content is better handled as a separate issue from instructional design, although well designed materials I have found are not weak on quality of content indicating that the design-content relationship is a synergistic one. The problem is not an unfamiliar one and is not unlike another issue, whether one stresses process or content. Unfortunately, in the past the issue was framed with process and content occupying polar positions and being irreconcilable antagonists, but now a more comprehensive view sees them as interactive variates of a successful educational program pointed toward production of specific competencies and behavioral patterns. Some further evidence from field testing does find, nevertheless, that content in some subject fields has taken precedence over concern about instructional design and systematic examination does reveal inconsistencies and contradictions in design. Two brief examples from social studies packages illustrate this point.

One social studies learning package written with a readability at several grade levels below its assigned grade level was developed for slow readers. Artfully designed it had programmed its lessons around a methodology which is well established, SQ3R. In a field assessment a group of teachers concluded that the methodology recommended was limited and contradictory to the intended objectives of the book. Relying heavily upon a methodology where the learner had to rely upon reading and verbal skills to cope with the learning package, the mode of transaction seemed inconsistent for the basic objective of a learning package for slow readers. Thus, they concluded that use of

this learning package in social studies would require considerable augmenting by the teacher of the recommended methodology.

A second social studies package was examined which relied heavily for the accomplishment of its objectives on the use of original historical documents and content analysis. Highly complex tasks on judging historical evidence were presented in the workbook without clearly identifying the necessary skills or making provision for their development. This is not an unusual defect and is common enough in workbooks which accompany textbooks that one wonders what type of planning goes into development of this part of the instructional package. Too frequently workbooks composed of a miscellany of tasks unrelated in sequence and bearing little or no defined relationship to any scope or sequence of the subject area. Accompanying evaluation instruments are equally informative on the producer's conception of design. In far too many cases I have found pronouncements on the importance of behavioral objectives are followed with fact-bound, stereotyped test items emphasizing content and of no assistance in evaluating pupil behavior.

For optimum results in using the instrument it is necessary to have a training period. Teachers are not used to looking at materials analytically and in the training period I have found old and cherished instructional packages take on new form after assessment. "I used to think this was a great series," is a frequent comment. But disallusionment is not the goal, rather a knowledge of the constraints and potential of the learning package - a critical awareness is the end sought. With several administrations teachers become quite adept in assessment; however, the first time through it is time consuming.

Are we at a stage where we can insist that instructional packages take explicit cognizance of some principles of instructional design and reflect some consistency in use of these principles? I have always thought judging a book by its cover had limitations - but I am inclined to believe that most of our assessment of instructional materials has not moved much beyond. The influence of instructional materials on curriculum and instruction has been noted by Jovanovich; "The schools inscribed a pattern, the publishers issued books to fit it, and in that gradual transmutation that became usual over the past half-century, the books made the course as often as the course made the books."[4] I would hope that the instrument I have presented here spawns criticism in producing, assessing, and utilizing instructional materials and assists in focusing the search for a more scientific and rational approach in these three areas.

Maurice J. Eash
University of Illinois
at Chicago Circle
March 1970

FOOTNOTES

[1]Development of this instrument was performed pursuant to a grant from the United States Office of Education to the New York State Network of Special Education Instructional Materials Centers, Hunter College of the City University of New York. Acknowledgement is extended for the assistance of Professor Gloria Wolinsky and her staff at various stages of the development and particularly for cooperation in arranging the initial field testing.

The author also wishes to acknowledge the invaluable assistance of Mr. Richard Smock and Mr. Edward Kelly in analyzing the data and their many helpful critical suggestions for improvement can be found in their Report #110, A Preliminary Instrument for Assessing Curriculum Materials: A Field Report and Discussion (mimeograph) Course Development Division, Office of Instructional Resources, University of Illinois, Urbana, Illinois. September 1969.

[2]Robinson, Helen et al. Calvalcades, Book 6, (Teachers Edition) Calvalcades, Book 6, (Student Edition) Think and Do, (Student Edition) Scott, Foresman and Company, 1965.

[3]Board of Education of the City of New York. Science Grade 7: The Chemistry of Matter, Curriculum Bulletin, #9a, 1962-63 Series. (Experimental Edition), 1963. 43 pp.

[4]Jovanovich, William. Now Barabbas. Harper. 1964. p. 56.

Maurice J. Eash
Professor of Education
University of Illinois
at Chicago Circle

An Instrument for the Assessment of Instructional Materials (Form IV)

I. OBJECTIVES	Yes	No
A. Are there objectives stated for the use of the material?	___	___
1. General objectives?	___	___
2. Instructional objectives?	___	___
3. Are the objectives stated in behavioral terms?[1]	___	___
4. If stated in behavioral terms, do the objectives specify:		
a. The type of behavior?	___	___
b. Conditions under which it will appear?	___	___
c. Level of performance expected?	___	___
5. List examples of objectives		

B. If there are no objectives stated for the use of the material, are the objectives instead implicit[2] or readily obvious? ___ ___

1. If yes, please outline below what objectives <u>you</u> believe govern the purpose of the material.

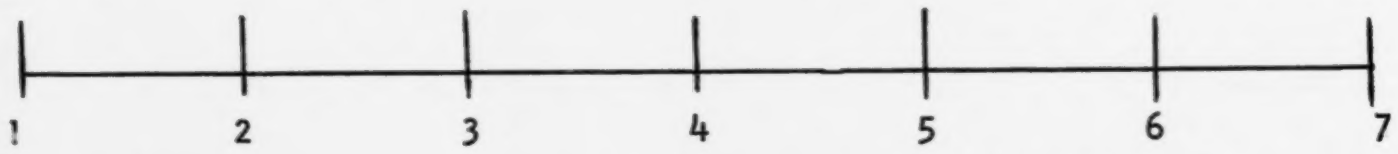
C. What appears to be the source of the objectives (both stated and implicit objectives)?

1. Are the objectives related to a larger frame of instruction? ___ ___
2. Are the objectives specific to a subject skill? ___ ___
3. Are the objectives related to a broader behavioral pattern[3] that is to be developed over a period of time? ___ ___
4. What seems to be the emphasis of the objectives: (Check as many as appropriate)
   a. Attitudinal[4] ___ b. Motor skills ___
   c. Cognitive development skills[5] ___ d. Subject skills ___
5. Are the objectives drawn from: (Check as many as appropriate)
   a. A learning approach[6] ___ b. Society needs (Citizenship) ___
   c. Demands of subject ___ d. Demands and needs of child[7] ___

Objectives continued

D. Quantitative rating of objectives

(DIRECTIONS: Please make an X on the rating scale below at the point which represents your best judgment on the following criteria. Please place the X ON a specific point.)



Objectives - vague, unclear, or missing. Those included not useful. Fails to distinguish between general and instructional objectives, mixes various types of objectives, confusing to the teacher.

Average, some of the criteria for objectives met, some missing, at times inconsistent, objectives only partially operational for the classroom teacher.

The objectives are stated clearly and in behavioral terms. Both general and instructional objectives are stated in a consistent conceptual framework. Excellent, one of the best, useful for a teacher.

II. ORGANIZATION OF THE MATERIAL (SCOPE AND SEQUENCE) — Yes — No

A. Has a task analysis[8] been made of the material and some relationship specified between the tasks? ___ ___

B. If a task analysis has been made, what basis was used to organize the materials: (Check as many as appropriate)

1. Errorless discrimination[9] ___
2. Simple to complex
3. Figure-ground[10] ___
4. General to specific
5. Logical order ___
6. Chronology ___

C. If no indication of a task analysis has been made, what assumptions do you believe the authors have made concerning the organization of the instructional sequence of the material?

D. Is there a basis for the scope of the material included in the instructional package? ___ ___

1. If there is a basis, is it:
   a. Related to a subject area ___ ___
   b. To a motor skill development ___ ___
   c. To a cognitive skill area ___ ___
   d. To an affective response system[11] ___ ___
   e. Other (specify) _____
2. Has the scope been subjected to analysis for:
   a. Appropriateness to students ___ ___
   b. Relationship to other material ___ ___

E. Is there a recommended sequence? ___ ___

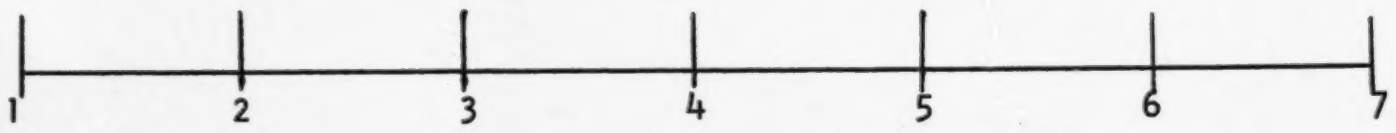
1. What is the basis of the recommended sequence? (Check as many as appropriate)
   a. Inter-relationships of a subject[12] ___
   b. Positive reinforcement and programed sequence[13] ___
   c. Open ended development of a generalization[14] ___
   d. Advanced organizer (cognitive)[15] ___
   e. Other (please specify) _____

F. Briefly outline the scope and sequence

Organization of Material continued

G. Quantitative rating of organization of the materials (Scope and Sequence)

(DIRECTIONS: Please make an X on the rating scale below at the point which represents your best judgment on the following criteria. Please place the X on a specific point.)



Sequence illogical or unstated, teacher is left to puzzle it out. Does not appear to have subjected material to any analysis to build an instructional design. Scope is uncertain, seems to contradict sequence. Little help unintentionally to teacher or children in organizing material.

Average in organization. Some help but teacher must supply much of organizational sequence. Scope somewhat limited, may be too narrow (or broad). Sequence is not detailed enough and may not have been tested with a range of children.

Excellent organization of scope and sequence. Conceptually developed based on a consistent theory; task analysis or other appropriate investigation has been done. Tested for appropriateness, of recommended sequence.

III. METHODOLOGY

	Yes	No
A. Does the author(s) and/or material suggest any methodological approach?	___	___
B. Is the methodological approach, if suggested, specific to the mode of transaction?	___	___
1. Does the mode of transaction[16]: (Check as many as appropriate)		
a. Rely upon teacher-centric method[17] (largely teacher directing?)	___	___
b. Rely upon pupil-centric method[18] (largely self-directing?)	___	___
c. Require active participation by the students?	___	___
d. Passive participation by the students?	___	___
e. Combination of active and passive participation by the students?	___	___
f. Direct students' attention to method of learning as well as the learning product?	___	___
g. Provide for variation among students - uses several approaches to method?	___	___
C. Does the methodology suggested require extensive preparation by the teacher?	___	___

1. How much deviation is permitted in methodology?

   Much ___ Some ___ Little ___

2. Does the methodology require unusual skills obtained through specific training?	___	___
3. Is there any statement on how methodology was tested: any experimental evidence?	___	___

4. If you have tried the recommended methodology, how successful did it seem for your students?

   Most succeeded ___ Approx. half succeeded ___ Few succeeded ___

   a. Please provide a brief description of the students who <u>were</u> successful and those who <u>were</u> <u>not</u> successful.

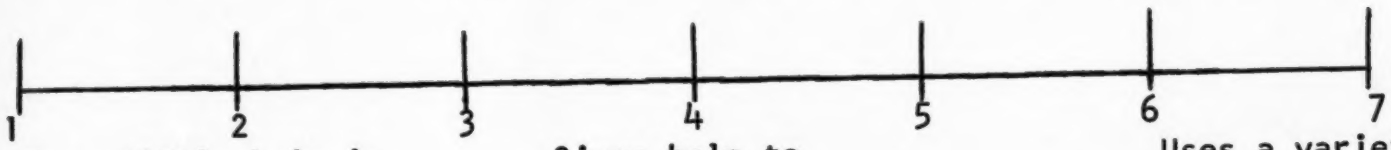
   b. What variations on recommended methodology have you used?

Methodology continued

D. In a brief statement describe the recommended methodology.

E. Quantitative rating of methodology.

(DIRECTIONS: Please make an X on the rating scale below at the point which represents your best judgment on the following criteria. Please place the X on a specific point.)



1 Very little help is given on methodology, or methodology is too abstract and complex for most students and teachers. Methodology appears to be unrelated to content and an afterthought in the learning package. Too active or passive for most students. Teacher required to participate fully with too many students at every step. Doesn't have appropriate methodology for variety of learning ability among students.

4 Gives help to the teacher, but would like more. Some students would be able to cope with suggested methodology, but others not. Doesn't appear to have been widely field tested. Teacher has to work out variety for students with special learning difficulties.

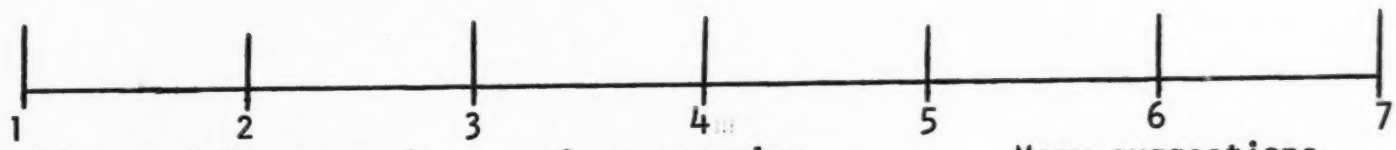
7 Uses a variety of modes in the transactions. Does not chain a teacher to a mode without reason, but provides assistance for different abilities. Describes the field test of the methodology. Teachers will find methodology easy to use and believe students will respond. Methodology is part of goals of instruction and not just vehicle for content.

IV. EVALUATION	Yes	No
A. Are there recommended evaluation procedures for teachers and students in the instructional package?	___	___
1. What do the evaluation procedures emphasize? (Check as many as appropriate)		
a. Cognitive skills ___ b. Subject skills ___		
c. Psychomotor skills[19] ___ d. Affective responses[20] ___		
2. Are the evaluation procedures compatible with the objectives?	___	___
3. Are evaluation procedures developed for several different levels: (Check as many as appropriate)		
a. Immediate feedback evaluation for the pupil ___		
b. Evaluation for a variety of the areas in #1 above, and over a period of time ___		
c. Immediate feedback evaluation for the teacher ___		
d. Evaluation on a norm referent[21] ___		
e. Evaluation on a criterion referent[22] ___		
B. Are the evaluation procedures contained in the package?	___	___
C. Does the evaluation give attention to both product and process learning?	___	___
D. Is there information on how evaluation procedures were tested and developed?	___	___
E. Briefly state what evaluation procedures are included. If possible, give examples.	___	___

Evaluation continued

F. Quantitative rating evaluation

(DIRECTIONS: Please make an X on the rating scale below at the point which represents your best judgment on the following criteria. Place the X on a specific point.)



Haphazard in approach. Product and process learnings either entirely neglected or confused. Lists items, but poorly constructed, no evidence of testing of evaluation approach. Students receive no assistance through feedback. Fails to recognize and examine different types of learning where appropriate.

Some examples given, range of evaluation limited. Samples given but limited and sketchy. Teacher finds useful that which is given, but needs more examples. Evaluation is limited to product or process. Unsure on whether evaluation has ever been tested, but seems logical though limited in types of learning examples.

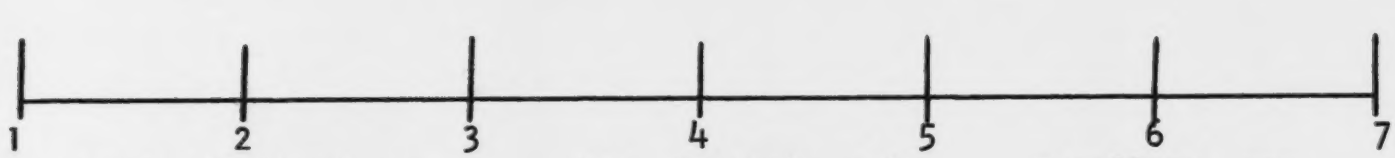
Many suggestions and helps in evaluation for the teacher. Has criterion reference procedures where appropriate. Student obtains assistance in learning through feedback evaluation. Gives attention to several kinds of learning, consistent with objectives of learning package.

COMMENT

A. Draw up an overall statement of the strengths and weaknesses of the material as an instructional package. Prepare your statement as if it were to be addressed to your fellow classroom teachers who are going to use it to make a decision on these instructional materials.

B. Quantitative rating overall assessment of material.

(DIRECTIONS: Please place an X on the point in the rating scale which best represents your overall judgment of these materials. Place the X on the specific point.)



Poorly designed, conceptually weak and inconsistent or haphazard design. Does not appear to have been field tested: inaccurate assumptions about children who will be using material. Overpriced, underdeveloped, a bad bargain.

Has strengths and weaknesses, but most teachers would find satisfactory, On the balance comes out about average, would need considerable supplementary effort by teacher. A compromise of price and availability.

Excellent, one of the best by comparison with other available material. Theoretically strong and carefully field tested. Shows consistent instructional design. Would recommend highly, well worth the price.

A GLOSSARY OF TERMS
USED IN THIS INSTRUMENT

1. Objectives stated in behavioral terms - a work picture of the type of behavior product which one might expect when the objective is achieved. Objectives stated in behavioral terms will usually name the behavior, state the conditions under which it will appear, and the level of performance expected, e.g. the child will be able to spell (type of behavior), in formal and informal writing (condition under which it will appear), 98 percent of the words in his written work (level of performance).

2. Implicit objectives - an examination of the content will permit the reader to readily identify the objectives that the student should accomplish, even if the producer has not stated them. If a filmstrip gives the sequential steps in solving arithmetic problems using long division, one would assume the implicit objective to be to teach the student the process of long division.

3. Broader behavioral pattern - instructional materials frequently are geared to goals that include complex behavior which is to be developed over time. Example: voting behavior as a function of citizenship involves a broader behavioral pattern which chains together a complex of behaviors ranging from knowing the candidates and the issues, to being registered, and knowing how to operate a voting machine. The instructional material may be designed to contribute to a broader behavioral pattern rather than a simpler, more specific behavior. Even if the objective is geared to a single specific behavior there should be some relationship to a broader behavioral pattern.

4. Attitudinal objectives - objectives that are designed to develop feelings and predispositions to act in accordance with internalized values and beliefs. These may be listed as attitudes, values, interests, and appreciations. They may be fairly direct as to develop in each student an interest in listening to a newscast at least once a day, or more complex as to form an attitude of critically evaluating the news by investigating the source of reports.

5. Cognitive development skills - objectives which have cognitive development skills (thinking) as a basis will usually emphasize thinking processes as their focus, such as understanding, discriminating, utilizing, chaining, and evaluating as opposed to emphasizing specific subject products.

6. Objectives drawn from a learning approach - objectives may be drawn utilizing approaches to learning, in some cases emphasizing wholeness of learnings prior to fragmenting into specifics for instruction. Example: the student will become familiar with the background of the 12th and 13th century European interest in colonies and trade, prior to studying the specific explorations. The extreme of the above approach would be a small step by step sequencing of the material on Europe in the 12th and 13th century in which concepts on European

8. Task analysis - the materials have been developed into specific tasks for the learner which have behavioral requirements that suggest a sequence for presentation and which allow an observer to determine if the learner accomplishes the task.

9. Errorless discrimination - the tasks are sequenced in such a manner that the student should move from step to step without making errors. This technique is used in some types of programmed instruction.

10. Figure-ground - the organization of materials, frequently perceptual in nature, in a field so that one stands out in a distinct way (figure) and the rest remains in the background (ground). Figure-ground organization can be used with other characteristics such as sounds, where one sound is heard over and above a background of others.

11. To an affective response system - where recognition is given to different levels of attitudes, from the simplest of merely attending to an object, to the building up of complex attitudes which predispose one's behavior toward a wide range of stimuli, e.g. enjoying a variety of forms of music.

12. Interrelationships of a subject - where the subject matter contains a logical relationship of concepts and processes. Example: adding must be mastered prior to multiplying. The local community is studied prior to more distant entities of state or federal government.

13. Positive reinforcement and programmed sequence - where the material has been developed into small steps that lead the learner toward a larger concept through a sequence that permits the learner to receive reinforcement through knowledge of right answers.

14. Open ended development of generalization - the instructional sequence is purposely quite open, e.g., letting the learner try out many possibilities and alternatives before arriving at a generalization.

15. Advanced organizers (cognitive) - a framework of key concepts, crucial to understanding and relating concepts of the larger body of material, are strategically placed in the sequence, forming an ideational ladder to which other material can readily be related. In some materials a short summary preceding the main body of instructional material delineates the key concepts or stresses their relationships to other concepts known by the learner, thus serving as advance organizers through the ideational anchors it gives to the learner for organizing, relating and remembering the new material.

16. Modes of transaction - a transaction is the interaction of a learner and stimuli in this context consisting of instructional materials. A mode is the channel that is used. Is the student asked to passively view, manipulate, verbally organize? Is the teacher an important part of the mode through exercising control over the learner's channels of transaction (methodological) to be used with instructional materials.

17. Teacher-centric method - the teacher is largely responsible for choosing and directing the mode of transaction for the learner. Teacher-centric modes of transaction prescribe that the "teacher will ..." and are predicated on obtaining specific learner responses.

18. Pupil-centric method - the learner is responsible for choosing the modes of transaction with the instructional material and is frequently left to evaluate and revise his behavior toward materials without teacher supervision.

19. Psychomotor skills - muscular or motor skills which require manipulation of material or objects. The ability to stack blocks is a psychomotor skill.

20. Affective response - responses which emphasize feelings, emotion or degree of acceptance or rejection stemming from internal attitudinal sets. Such responses may be labelled attitudes, biases, interests, etc.

21. Norm referent evaluation - judging a learner's performance by what other known groups of learners do on the same tasks. Achievement test scores, aptitude tests and mental test scores report their results in norm referent terms. The statement, "This particular learner scored at 4th grade level," is using a norm referent evaluation of the learner's performance.

22. Criterion referent evaluation - the learner is judged on his ability to do a specified task or demonstrate the behavior appropriate to the task. The learner is judged on whether he can or cannot demonstrate the appropriate behavior that signifies task accomplishment and is not judged by comparison of his performance with another group of learners.